REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
29-02-2012	Final	15-03-2009 to 29-02-2012

4. TITLE AND SUBTITLE
Causal Networks with Selectively Influenced Components

5a. CONTRACT NUMBER

5b. GRANT NUMBER
FA9550-09-01-0252

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Schweickert, R.
Dzhafarov, E. N.
Kujala, Janne V.

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Purdue University
Dept. of Psychological Sciences
703 Third St.
West Lafayette, IN 47907

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Office of Scientific Research
875 N. Randolf St., Room 3112
Arlington, VA 22203

10. SPONSOR/MONITOR'S ACRONYM(S)
AFOSR

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AFRL-OSR-VA-TR-2012-0520

12. DISTRIBUTION / AVAILABILITY STATEMENT
Distribution A--Approved for public release

13. SUPPLEMENTARY NOTES
None

14. ABSTRACT
In work superivised by Dr. Schweickert, a method for inferring a processing tree to account for response probability was generalized to allow for multiple response classes, parameters such as rates not bounded above by 1, and experimental factors having effects at more than one vertex. A further generalization allows a processing tree to be inferred from joint analysis of reaction time and accuracy. Processing trees were inferred for immediate serial recall data and data on sleep and interference. In the part under Dr. Dzhafarov's supervision, a general mathematical theory of selective influences was elaborated (which input influences which of probabilistically interdependent random outputs); the Joint Distribution Criterion was formulated in complete generality; a theory of pseudo-quasi-metrics was constructed to be used to test for selectiveness of influences; a Linear Feasibility Test for selective influences with finite-valued random outputs was constructed; a formal equivalence of selective influences with the issue of quantum entanglement in physics was established, with non-commuting measurements in quantum physics paralleling the mutually exclusive values of inputs (external factors) in behavioral sciences.

15. SUBJECT TERMS
critical path network, joint distribution criterion, linear feasibility test, processing tree, pseudo-quasi-metric tests, selective influences, quantum entanglement

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Richard Schweickert
U	U	U	UU	22	**19b. TELEPHONE NUMBER** *(include area code)* 765 494 7986

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

This report consists of two parts. The first summarizes work done under supervision of Dr. Schweickert. The second (beginning after p. 6 of the first part) summarizes work done under supervision of Dr. Dzhafarov

Causal networks with selectively influenced components

Report on the part under Richard Schweickert's supervision

Considerable evidence indicates that the mental processes involved in performing some tasks, such as recall of items from a list, are arranged in a processing tree (Batchelder & Riefer, 1999). In such a tree, each process is represented by a vertex and each possible outcome of a process is represented by an arc descending from the vertex representing it. Many causal networks can be represented as processing trees (Schweickert & Chen, July, 2008). Responses are represented by terminal vertices that have no descendents. Processing trees are usually used to model the probabilities of the various possible responses. Typically an investigator proposes such a tree, estimates parameters, and tests it through goodness of fit. A portion of the work on the grant is on developing further another approach, Tree Inference (Schweickert & Chen, 2008). With Tree Inference, a processing tree is not proposed ahead of time. Instead, the investigator manipulates experimental factors, such as the number of items to be recalled and the delay between presentation and recall. A factor is said to selectively influence a vertex if it changes parameters associated with the descendents of that vertex and no other. If a factor selectively influences a vertex we also say it selectively influences the process represented by that vertex. In an experiment with two factors, the investigator can test whether each factor selectively influences a different vertex. If so, the form of a processing tree accounting for the data can determined.

Prior to the work on the grant, processing trees were not used for modeling reaction time, and there were three limitations to Tree Inference. 1) It was applicable to experiments with only two possible responses (e.g., correct or wrong). 2) Parameters associated with the arcs of a processing tree were probabilities bounded above by 1. 3) An experimental factor was required to have an effect at only one vertex. These three limitations have now been overcome (Schweickert & Xi, 2011). Now, for example, rates of responding can be analyzed in addition to probabilities of responses.

A pair of vertices can be related in one of two ways in a processing tree. There may be a path from the root to a terminal vertex that passes through both vertices. In that case the vertices are said to be ordered. If there is no such path, the vertices are said to be unordered. Qualitative tests have been developed, through work on the grant, that allow an investigator to test whether two experimental factors selectively influence two ordered vertices, and if so, determine their order. Processing trees were found to account well for data in the literature on immediate ordered recall and on effects of sleep and retroactive interference (Schweickert, Fisher & Sung, in press).

Processing trees were originally developed for analyzing response probabilities, not response times. Through recent work on the grant, processing trees can now be inferred from a joint analysis of response time and accuracy. New theorems provide necessary and sufficient conditions for reaction time data to be generated from an experiment in which two factors selectively influence two different processes.

Although processing trees describe well the organization of mental processes for some tasks, there is no reason to expect processes to be organized in the same way for all tasks. In some tasks, there is evidence that the processes are organized in a directed acyclic network (a critical path network). In earlier work (Schweickert, 1978), a method was developed for analyzing reaction times to test whether two experimental factors selectively influence two different processes in a directed acyclic network. If the test was passed, part of the network could be inferred from the data. In particular, an investigator could determine whether the two selectively influenced processes are sequential (ordered in the network) or concurrent (unordered).

Ordinarily for a given data set, if one directed acyclic network can account for the data, then several different networks can account for the data as well. There are two ways to determine the form of a directed acyclic network that accounts for the reaction time data when factors selectively influence processes in the network. One way is quantitative, through analysis of the slacks in the network (Schweickert, 1978). Another way is qualitative, through analysis of which pairs of processes are sequential and which are concurrent, using the Transitive Orientation Algorithm (e.g., Golumbic, 1980). Each method generates a set of possible networks that can account for the data, so the question arises of whether the set of possible networks generated by one method is more restricted than the set generated by the other. Work on the grant that shows that these sets are the same when a serial-parallel network accounts for the data. In other words, the uniqueness of directed acyclic networks inferred from the effects on reaction time of factors selectively influencing processes has now been characterized for serial-parallel networks. A case remains to be characterized, that of networks containing a subnetwork in the form of a Wheatstone bridge.

One of the products of the grant is a book, planned for appearance in early 2012, on inferring cognitive architecture by selectively influencing mental processes.

A manuscript related to the grant is now under review (Schweickert, Fortin, Xi & Viau-Quesnel). Because the data are not publically available, they are summarized in the Appendix.

References to work not arising from the grant

Batchelder, W. H., & Riefer, D. M. (1999). Theoretical and empirical review of multinomial process tree modeling. *Psychonomic Bulletin & Review, 6*, 57-86.

Golumbic, M. C. (1980). *Algorithmic graph theory and perfect graphs*. NY: Academic Press.

Schweickert, R. (1978). A critical path generalization of the additive factor method: Analysis of a Stroop task. *Journal of Mathematical Psychology, 18*, 105-139.

Schweickert, R., & Chen, S. (July, 2008). Processing trees equivalent to mixtures of processing trees. Society for Mathematical Psychology Meeting, Washington, DC, July.

Schweickert, R., & Chen, S. (2008). Tree inference with factors selectively influencing processes in a processing tree. *Journal of Mathematical Psychology, 52,* 158-183.

Publications

Peer-reviewed

Fortin, C., Schweickert, R., Gaudreault, R., & Viau-Quesnel, C. (2010). Timing is affected by demands in memory search, but not by task switching. *Journal of Experimental Psychology: Human Perception and Performance, 36,* 580-595.

Schweickert, R., Fisher, D. L., & Goldstein, W. M. (2010). Additive factors and stages of mental processes in task networks. *Journal of Mathematical Psychology, 54,* 405-414.

Schweickert, R., & Xi, Z. (2011). Multiplicatively Interacting Factors Selectively Influencing Processes in Multiple Response Class Processing and Rate Trees. *Journal of Mathematical Psychology,* 55, 348-364.

Fisher, D. L., Schweickert, R., Drury, C. (In press). Mathematical models in engineering psychology: Optimizing performance. In G. Salvendy (Ed.), *Handbook of human factors and ergonomics*, 4 th ed. Hoboken, NJ: Wiley (pp 997-1024).

Not Peer-reviewed

Han, H. J., Viau-Quesnel, C., Xi, Z., Schweickert, R., & Fortin, C. (2010). Self-Timing in memory and visual search tasks. In Bastianelli, Alessia, & Vidotto, Giulio (Eds.) Fechner Day 2010. *Proceeding of the 26th Annual Meeting of the International Society for Psychophysics,* Padua, Italy: The International Society for Psychophysics, 363-368.

Schweickert, R. (In press). Network models of tasks. In J. D. Wright (Ed.), *International Encyclopedia of the Social and Behavioral Sciences,* 2 nd Ed. Oxford: Elsevier.

Book

Schweickert, R., Fisher, D. L., & Sung, K. (In press). *Discovering cognitive architecture by selectively influencing mental processes.* Singapore: World Scientific.

Manuscript submitted for publication

Schweickert, R., Fortin, C., Xi, Z., & Viau-Quesnel, C. (Submitted). Memory set search affects concurrent timing, memory set activation does not. Submitted for publication, *Journal of Experimental Psychology: Human Perception and Performance.*

Conference Presentations

Schweickert, R., Fisher, D. L., & Goldstein, W. L. (2009). Additive factors and stages of mental processes. Meeting of Society for Mathematical Psychology and European Mathematical Psychology Group, Amsterdam, August.

Schweickert, R., & Xi, Z. (2009). Multiplicatively interacting factors in multiple response class processing trees. Meeting of Society for Mathematical Psychology and European Mathematical Psychology Group, Amsterdam, August.

Schweickert, R., & Xi, Z. (2009). Human Associative Memory and Sleep: The Dream World is a Small World. Poster presented at the Complexity Conference, Institute on Complex Systems, Northwestern University, September.

Schweickert, R., Batchelder, W. H., & Xi, Z. (2009) Organizers, Winer Lectures on Multinomial Trees and Related Models. Purdue University, October.

Schweickert, R., & Xi, Z. (2009). Tree Inference Through Selective Influence of Factors in Processing Trees. Paper presented at Winer Lectures, Purdue University, October.

Schweickert, R. (2010). Selective Influence in the Ratcliff Diffusion Model. AFOSR Joint Program Review--Cognition and Decision Program and Human-System Interface Program, Arlington, VA, January.

Viau-Quesnel, C., Schweickert, R., & Fortin, C. (2010). L'étude des réseaux sociaux en psychologie cognitive : un exemple basé sur l'étude de l'organisation des rêves. Société Québécoise pour la Recherche en Psychologie. Montreal, April.

Schweickert, R., Viau-Quesnel, C., Xi, Z., Han, H. J., & Fortin, C. (2010). Regularities in dream social networks. Portland, OR, August.

Viau-Quesnel, C., Fortin, C., & Schweickert, R. (2010). Task switch costs on concurrent timing: an apparent dissociation between voluntary and involuntary task switching. Meeting of the Psychonomic Society, St. Louis, November.

Schweickert, R. (2011). Determining the order of mental processes. AFOSR Program Review--Mathematical Modeling of Cognition and Decision Program, Dayton, January.

Xi, Z., & Schweickert, R. (2011). Tree inference with multiplicatively interacting factors in a multiple response processing tree. Midwest Cog Science Meeting, East Lancing, MI April.

Han, H. J., Viau-Quesnel, C., Xi, Z., & Schweickert, R. (2011). Self-timing in memory and visual search. Midwest Cog Science Meeting, East Lancing, MI April.

Schweickert, R. (2011). Chair and Organizer, Symposium on Memory Consolidation and Structure in Sleep and Dreams.   Annual Convention of the Association for Psychological Science, Washington, DC, May.

Schweickert, R. (2011). Determining the order of mental processes by selectively influencing them. Meeting of the Society for Mathematical Psychology, Boston, July.

Schweickert, R., Xi, Z., Viau-Quesnel, C., Yamaguchi, M., & Fortin, C. (2011). Activating a memory set and searching it have different effects on time production. Psychonomic Society Meeting, Seattle, November.

Appendix

Manuscript of Schweickert, Fortin, Xi & Viau-Quesnel (Submitted): Method and Results in Brief

Method. In the experiment summarized here, the participant began by memorizing two memory sets, a set of words and a set of consonants. One memory set was presented again at the start of each trial. Items in this memory set were said to be from the active pool.

In the Reaction Time Condition, on each trial, a probe was presented and the participant's task was to press a button to indicate whether the probe was present in one of the memory sets or not. Two factors were varied from trial to trial: the presence or absence of the probe in the memory set and whether the probe was from the active memory pool or the inactive memory pool. A third factor was varied between blocks of trials, whether the active and inactive memory set contained three and six items, respectively, or contained six and three items, respectively. Reaction time and accuracy were measured.

In the Time Production Condition, the procedure was the same, except as follows. Prior to the memory search trials, participants were trained to produce a time interval of 2.4 seconds. On the later memory search trials, the participant was instructed to respond when he or she judged that a 2.4 second interval had elapsed since presentation of the probe.

Results. There were two main results. First, in the Reaction Time Condition, reaction times increased when the inactive memory set was searched and increased when the set size was larger. The combined effect of these two factors was additive. The probability of a correct response decreased when the inactive memory set was searched and decreased when the set size was larger. The combined effect of these two factors was multiplicative. A simple processing tree accounts for the data well, in which activation of the memory set is followed by searching the memory set. The durations of the processes add and the probabilities that the processes are correct multiply.

The second main result is in the Time Production Condition. Time intervals produced by the participants were longer when the size of the memory set to be searched was larger. The memory search interfered with timing. However, the time intervals produced were not longer when the memory set was inactive. Activating the memory set did not interfere with timing. Timing and activating a memory set do not compete for capacity, but timing and search do.

Table from Schweickert, Fortin, Xi & Viau-Quesnel (Submitted)

Mean Reaction Times, Time Productions and Percent Errors

Reaction Time Condition

	Memory Set			
Probe	Active Size 3	Active Size 6	Inactive Size 3	Inactive Size 6
Present	830 (3.5)	933 (6.2)	921 (10.7)	1003 (12.6)
Absent	880 (1.1)	969 (2.8)	928 (3.7)	1016 (7.0)

Time Production Condition

	Memory Set			
Probe	Active Size 3	Active Size 6	Inactive Size 3	Inactive Size 6
Present	3419 (3.6)	3530 (8.0)	3530 (13.9)	3463 (13.7)
Absent	3418 (1.8)	3548 (4.4)	3530 (7.0)	3469 (10.9)

Note: Times in msec, percent errors in parentheses.

Causal networks with selectively influenced components

Report on the part under Ehtibar Dzhafarov's supervision

1 Introduction

1.1 The problem and its applications

This part of the project was aimed at further developing the theory of selective probabilistic causality. The theory answers the question: Given a set of inputs into a system (e.g., independent variables characterizing stimuli in a psychological experiment) and a set of stochastically non-independent random outputs (e.g., random variables describing different aspects of human responses), how can one determine, for each of the outputs, which of the inputs it is influenced by?

The theory has applications in behavioral and social sciences, including such problems as: in the investigations of networks of mental operations, does a certain experimental manipulation selectively influence only a certain component of the network? in conjoint testing, does study time or specific training for one of the tests selectively influence one's performance in this test only? in studying perceptual judgments, is an assessment of a given stimulus property selectively influenced by this property alone? in medical research, does the presence or absence of a given symptom selectively depend on a given illness?

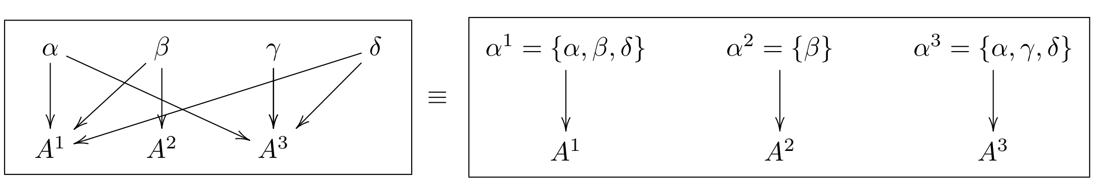
The theory also has applications in quantum mechanics, in answering such questions as: can a model with local non-contextual variables account for the distribution of spins in a system of entangled particles? The non-commuting measurements of spins along different axes performed on a given particle correspond to the mutually exclusive values of an experimental manipulation in behavioral applications.

Other applications of the theory can be deduced from the fact that it generalizes all conceivable combinations of nonlinear factor and regression analyses, with no constraints imposed on the relationship between explanatory and response variables, or on the unobservable sources of randomness.

This part of the project involved as a senior personnel Janne Kujala of University of Jyväskylä, Finland.

1.2 Notation and basic definitions

The problem can be illustrated on the following *diagram of selective influences*, shown in two equivalent forms:



A^1, A^2, and A^3 here are *random outputs*, α, β, γ, and δ are *inputs* (also referred to as *external factors*), and arrows indicate the relation "influences." The right-hand diagram is in the *canonical form*, with the factors redefined so that each random output A^i is influenced by a single factor, α^i. Factors are treated as deterministic quantities, i.e., even if they are random variables, the joint distribution of the outputs is always conditioned on their specific values (or *levels*). Each factor can be on one of several *levels*, and the joint distribution of $\left(A^1, A^2, A^3\right)$ is supposed to be known for each allowable combination of factor levels (*treatment*). Thus, if factors $\alpha, \beta, \gamma, \delta$ are all binary, then each of the $2 \times 2 \times 2 \times 2 = 16$ logically possible combinations is a potential treatment, but the joint distribution of $\left(A^1, A^2, A^3\right)$ may only be considered for some of them (e.g., treatments that have not been used in the experiment or are physically impossible are not allowable). This is an important consideration if one wishes to conveniently deal with the canonical diagrams of selective influences only. In our example, $\alpha^1, \alpha^2, \alpha^3$ have 8,2, and 8 levels, respectively, but the number of allowable treatments cannot exceed $16 < 8 \times 2 \times 8$.

The general theory and the pseudo-quasi-metric tests discussed in Section 2.2 have been developed for arbitrary sets of factors and outputs, but to keep notation simple this report is confined to finite sets only. Also, for simplicity only, the random outputs are assumed to be random variables in the narrow sense (corresponding to the conventional Lebesgue measure or countable measure); they may be arbitrary in the general theory.

Let Λ_i be the set of possible levels of factor α^i $(i = 1, \ldots, n)$ in a canonical diagram of selective influences

$$\boxed{\begin{array}{ccccc} \alpha^1 & \ldots & \alpha^i & \ldots & \alpha^n \\ \downarrow & & \downarrow & & \downarrow \\ A^1 & \ldots & A^i & \ldots & A^n \end{array}}$$

Let $\Phi \subset \Lambda_1 \times \ldots \times \Lambda_n$ be the set of allowable treatments, and let for every treatment $\phi \in \Phi$ the joint distribution of $\left(A^1, \ldots, A^n\right)$ be given. The first question is:

> If at least for some of the treatments ϕ the random outputs are not stochastically independent, what is the meaning of saying that A^1 is *selectively* (exclusively) influenced by α^1, A^2 by α^2, etc.?

And assuming that a reasonable definition of selective influences is achieved (which means a definition satisfying certain desiderata, listed below, and lending itself to fruitful mathematical development), the second question is:

> How can one determine, based on the joint distributions of $\left(A^1_\phi, \ldots, A^n_\phi\right)$ for each $\phi \in \Phi$, whether the canonical diagram of selective influences is satisfied?

The reasonable definition in question can be given in one of two equivalent forms:

(SI_1) there are independent random entities $C, S^1, \ldots, S^n$ and functions $R^i\left(\alpha^i, C, S^i\right)$ $(i = 1, \ldots, n)$ such that

$$\left(R_1\left(j_1, C, S^1\right), \ldots, R_n\left(j_n, C, S^1\right)\right) \sim \left(A^1_\phi, \ldots, A^n_\phi\right)$$

for any $\phi = (j_1, \ldots, j_n) \in \Phi$ ($\sim$ meaning "identically distributed");

(SI_2) there is a random entity C and functions $R^i\left(\alpha^i, C\right)$ $(i = 1, \ldots, n)$ such that

$$\left(R^1\left(j_1, C\right), \ldots, R^n\left(j_n, C\right)\right) \sim \left(A^1_\phi, \ldots, A^n_\phi\right)$$

for any $\phi = (j_1, \ldots, j_n) \in \Phi$.

In quantum physics (see Section 2.4) these formulations correspond to the classical explanations of the entanglement phenomena with, respectively, stochastic and deterministic hidden variables.

The fact that either of these definitions (SI_1 or SI_2) is satisfied is schematically indicated as $\left(A^1, \ldots, A^n\right) \looparrowleft \left(\alpha^1, \ldots, \alpha^n\right)$.

2 Progress

2.1 Joint Distribution Criterion

Definitions SI_1 and SI_2 were shown to be equivalent to the following proposition, called the *Joint Distribution Criterion*:

(JDC) there is a jointly distributed vector of random variables

$$H = \left(H^1_1, \ldots, H^1_{k_1}, \ldots, H^i_1, \ldots, H^i_{k_i}, \ldots, H^n_1, \ldots, H^n_{k_n}\right)$$

such that

$$\left(H^1_{j_1}, \ldots, H^n_{j_n}\right) \sim \left(A^1_\phi, \ldots, A^n_\phi\right)$$

for any $\phi = (j_1, \ldots, j_n) \in \Phi$.

H is called the *JDC-vector*. This criterion has a greater heuristic power than definitions SI_1 and SI_2. Some of the immediate consequences of JDC are as follows:

1. for any subset $\{i_1, \ldots, i_k\}$ of $\{1, \ldots, n\}$, $\left(A^{i_1}, \ldots, A^{i_k}\right)$ does not depend on factors outside $\left(\alpha^{i_1}, \ldots, \alpha^{i_k}\right)$ (*complete marginal selectivity*);
2. for any subset $\{i_1, \ldots, i_k\}$ of $\{1, \ldots, n\}$ we have $\left(A^{i_1}, \ldots, A^{i_k}\right) \looparrowleft \left(\alpha^{i_1}, \ldots, \alpha^{i_k}\right)$ (*nestedness*);
3. for any measurable functions $F_1\left(\alpha^1, a^1\right), \ldots, F_n\left(\alpha^n, a^n\right)$ we have

   $$\left(F_1\left(\alpha^1, A^1\right), \ldots, F_n\left(\alpha^n, A^n\right)\right) \looparrowleft \left(\alpha^1, \ldots, \alpha^n\right)$$

   (*invariance with respect to factor-level-specific transformations* of random outputs);
4. if $\left(A^1, \ldots, A^n\right)$ are random variables in the narrow sense, then C in SI_2 or $C, S^1, \ldots, S^n$ in SI_1 can always be chosen to be random variables in the narrow sense. Moreover, they can be chosen arbitrarily as any continuously (atomlessly) distributed random variables, e.g., uniformly distributed between 0 and 1.

2.2 Distance-like functions

Let $X = \{X_\omega : \omega \in \Omega\}$ be an indexed set of jointly distributed random variables X_ω with distributions $(V_\omega, \Sigma_\omega, \mu_\omega)$. For any $\alpha, \beta \in \Omega$, the ordered pair (X_α, X_β) is a random variable with distribution $(V_\alpha \times V_\beta, \Sigma_\alpha \times \Sigma_\beta, \mu_{\alpha,\beta})$, and $X \times X$ is a set of jointly distributed random variables (hence also a random variable).

We call a function $d : X \times X \to \mathbb{R}$ a *pseudo-quasi-metric* (*p.q.-metric*) on X if, for all $\alpha, \beta, \gamma \in \Omega$,

(i) $d(X_\alpha, X_\beta)$ only depends on the joint distribution of (X_α, X_β),

(ii) $d(X_\alpha, X_\beta) \geq 0$,

(iii) $d(X_\alpha, X_\alpha) = 0$,

(iv) $d(X_\alpha, X_\gamma) \leq d(X_\alpha, X_\beta) + d(X_\beta, X_\gamma)$.

Conventional pseudometrics (also called semimetrics) obtain by adding the property $d(X_\alpha, X_\beta) = d(X_\beta, X_\alpha)$; conventional quasimetrics are obtained by adding the property $\alpha \neq \beta \Rightarrow d(X_\alpha, X_\beta) > 0$. A conventional metric is both a pseudometric and a quasimetric.

The relevance of the p.q.-metrics on the sets of jointly distributed random variables to the problem of selectivity lies in the general test (necessary condition) for selectivity of influences, formulated after the following definition.

We call a sequence of input points $(j_1, \ldots, j_l)$ (where $j_i \in \Lambda_i$ for $i = 1, \ldots, l \geq 3$) *treatment-realizable* if there are treatments $\phi^1, \ldots, \phi^l \in \Phi$ (not necessarily pairwise distinct), such that

$$(j_1, j_l) \subset \phi^1 \text{ and } (j_{i-1}, j_i) \subset \phi^i \text{ for } i = 2, \ldots, l.$$

Now, if a JDC-vector H exists, then for any p.q.-metric d on H we should have

$$d\left(H^1_{j_1}, H^l_{j_l}\right) = d\left(A^1_{\phi^1}, A^l_{\phi^1}\right)$$

and

$$d\left(H^{i-1}_{j_{i-1}}, H^i_{j_i}\right) = d\left(A^{i-1}_{\phi^i}, A^i_{\phi^i}\right)$$

for $i = 2, \ldots, l$ whence

$$d\left(A^1_{\phi^1}, A^l_{\phi^1}\right) \leq \sum_{i=2}^{l} d\left(A^{i-1}_{\phi^i}, A^i_{\phi^i}\right). \tag{1}$$

This chain inequality, written entirely in terms of observable probabilities, is referred to as a *p.q.-metric test* for selectivity of influences. If this inequality is violated for at least one treatment-realizable sequence of input points, no JDC-vector H exists, and the selectivity is ruled out.

It turns out that one only needs to check the chain inequality for a special subset of all possible treatment-realizable sequences $j_1, \ldots, j_l$. Namely, a treatment-realizable sequence $j_1, \ldots, j_l$ is called *irreducible* if $j_1 \neq j_l$ and the only subsequences $\{j_{i_1}, \ldots, j_{i_k}\}$ with $k > 1$ that are subsets of treatments are pairs $\{j_1, j_l\}$ and $\{j_{i-1}, j_i\}$, for $i = 2, \ldots, l$. Otherwise the sequence is *reducible*. It was proved that

> given a p.q.-metric d on the hypothetical JDC-vector H, inequality (1) is satisfied for all treatment-realizable sequences if and only if this inequality holds for all irreducible sequences.

As a special case,

if $\Phi = \Lambda_1 \times \ldots \times \Lambda_n$ (a completely crossed factorial design), then inequality (1) is satisfied for all treatment-realizable sequences if and only if this inequality holds for all tetradic sequences of the form x, y, s, t, with $x, s \in \{\alpha\} \times \Lambda_i$, $y, t \in \Lambda_j$, $x \neq s$, $y \neq t$, $i \neq j$.

A very versatile and useful class of p.q.-metrics is formed by *order-distances*. Given an indexed set of jointly distributed random variables $X = \{X_\omega : \omega \in \Omega\}$, let

$$R \subset \bigcup_{(\alpha,\beta)\in\Omega\times\Omega} V_\alpha \times V_\beta,$$

where V_ω denotes the set of possible values of X_ω. We write $a \preceq b$ to designate $(a, b) \in R$. Let R be a total order, that is, transitive, reflexive, and connected in the sense that for any $(a, b) \in \bigcup_{(\alpha,\beta)\in\Omega\times\Omega} V_\alpha \times V_\beta$, at least one of the relations $a \preceq b$ and $b \preceq a$ holds. We define the equivalence $a \sim b$ and strict order $a \prec b$ induced by $\preceq$ in the usual way. Finally, we assume that for any $(\alpha, \beta) \in \Omega \times \Omega$, the sets

$$\{(a, b) : a \in V_\alpha, b \in V_\beta, a \preceq b\}$$

are $\mu_{\alpha,\beta}$-measurable, where $\mu_{\alpha,\beta}$ is the probability measure for (X_α, X_β). This implies the $\mu_{\alpha,\beta}$-measurability of the sets

$$\{(a, b) : a \in V_\alpha, b \in V_\beta, a \prec b\}, \ \{(a, b) : a \in V_\alpha, b \in V_\beta, a \sim b\}.$$

The function

$$\mathrm{D}(X_\alpha, X_\beta) = \Pr[X_\alpha \prec X_\beta]$$

is called an *order p.q.-metric*, or *order-distance*, on X. It was proved that D satisfies the properties (i)-(iv) of the definition of a p.q.-metric.

As an example of an order-distance applied to the selectivity problem, let $\Lambda_1 = [0, 1]$, $\Lambda_2 = [0, 1]$, $\Phi = \Lambda_1 \times \Lambda_2$. Let A^1_ϕ, A^2_ϕ for any treatment $\phi = (w_1, w_2)$ have a bivariate normal distribution with zero means, unit variances, and correlation $\rho = \min(1, w_1 + w_2)$. Marginal selectivity is trivially satisfied. For any bivariate normally distributed (A, B), let us define $A \prec B$ iff $A < 0, B \geq 0$. Then the corresponding order-distance on the hypothetical JDC-set H is

$$\mathrm{D}\left(H^1_{w_1}, H^2_{w_2}\right) = \frac{\arccos(\min(1, w_1 + w_2))}{2\pi}.$$

The sequence of input points $(1, 0), (2, 1), (1, 1), (2, 0)$ is treatment-realizable, so if H exists, we should have

$$\mathrm{D}\left(H^1_0, H^2_0\right) \leq \mathrm{D}\left(H^1_0, H^2_1\right) + \mathrm{D}\left(H^2_1, H^1_1\right) + \mathrm{D}\left(H^1_1, H^2_0\right).$$

The numerical substitutions yield, however,

$$\frac{1}{4} \leq 0 + 0 + 0,$$

and as this is false, the hypothesis of selectively is rejected.

This example generalizes into a special class of order-distances, *classification distances*, defined by the following construction of $\preceq$: provided the sigma-algebra Σ_ω associated with each V_ω contains at least $n > 1$ disjoint nonempty sets, one can partition each V_ω as $\bigcup_{k=1}^n V^{(k)}_\omega$, with $V^{(k)}_\omega \in \Sigma_\omega$, and put $a \preceq b$ if and only if $a \in V^{(k)}_\alpha, b \in V^{(l)}_\beta$ and $k \leq l$. Another application of classification distances will be given in Section 2.4.

2.3 Linear Feasibility Test

Let now each random variable A^i have a finite set of m_i possible values (enumerated for simplicity $1, \ldots, m_i$), and let each factor/input α^i contain k_i factor levels (enumerated $1, \ldots, k_i$). This is arguably the most important special case both because it is ubiquitous and because in all other cases random variables and factors can be discretized into finite number of categories. The *Linear Feasibility Test* (LFT) to be described is a direct application of JDC to this situation, furnishing a necessary and sufficient condition for the diagram of selective influences $\left(A^1, \ldots, A^n\right) \looparrowleft \left(\alpha^1, \ldots, \alpha^n\right)$.

The distributions of $\left(A^1_\phi, \ldots, A^n_\phi\right)$ are represented by probabilities

$$\Pr\left[A^1_\phi = a_1, \ldots, A^n_\phi = a_n\right],$$

for all $\phi = (j_1, \ldots, j_n) \in \Phi$ and all $(a_1, \ldots, a_n) \in \{1, \ldots, m_1\} \times \ldots \times \{1, \ldots, m_n\}$. We consider this probability the $[(a_1, \ldots, a_n), (j_1, \ldots, j_n)]$th component of the $m_1 \ldots m_n t$-vector P of all such probabilities (with t denoting the number of treatments in Φ). The joint distribution of H in JDC, if it exists, is represented by probabilities

$$\Pr\left[H^1_1 = h^1_1 \ldots, H^1_{k_1} = h^1_{k_1}, \ldots, H^n_1 = h^n_1, \ldots, H^n_{k_n} = h^n_{k_n}\right],$$

with

$$\left(h^1_1, \ldots, h^1_{k_1}, \ldots, h^n_1, \ldots, h^n_{k_n}\right) \in \{1, \ldots, m_1\}^{k_1} \times \ldots \times \{1, \ldots, m_n\}^{k_n}.$$

We consider this probability the $\left(h^1_1, \ldots, h^1_{k_1}, \ldots, h^n_1, \ldots, h^n_{k_n}\right)$th component of the $(m_1)^{k_1} \ldots (m_n)^{k_n}$-vector Q of all such hypothetical probabilities.

Consider now the Boolean matrix M with rows corresponding to components of P and columns to components of Q: let $M(r, c) = 1$ if and only if row r corresponds to the $[(j_1, \ldots, j_n), (a_1, \ldots, a_n)]$th component of P, column c to the $\left(h^1_1, \ldots, h^1_{k_1}, \ldots, h^n_1, \ldots, h^n_{k_n}\right)$th component of Q, and

$$h^1_{j_1} = a_1, \ldots, h^n_{j_n} = a_n.$$

Clearly, the vector Q exists if and only if the system

$$MQ = P,\ Q \geq 0 \tag{2}$$

has a solution (is *feasible*). This is a linear programming task in the standard form (with a constant objective function). Let $\mathcal{L}(P)$ be a Boolean function equal to 1 if and only if this system is feasible. $\mathcal{L}(P)$ is known in linear programming to always be computable, its time complexity being polynomial. It is therefore justifiable to call JDC a general solution for the problem of rejecting or confirming a diagram of selective influences in all cases involving only finite sets of values/levels.

The linear system (2) is feasible if and only if the point P belongs to the convex hull of the points corresponding to the columns of M, which form a subset of the vertices of a unit hypercube. In particular, if the set Φ of allowable treatments contains all combinations of factors points, the polytope is the $((k_1(m_1 - 1) + 1) \ldots (k_n(m_n - 1) + 1) - 1)$-dimensional convex hull of the points corresponding to the columns of the Boolean matrix M, which form a subset of the vertices of the $(m_1)^{k_1} \ldots (m_n)^{k_n}$-dimensional unit hypercube.

As an example, let there be factors $\alpha = \{1, 2\}$, $\beta = \{1, 2\}$, and let the set of allowable treatments Φ consist of all four possible combinations of the factor points. Let A and B be binary variables, $a_1 = b_1 = 1$, $a_2 = b_2 = 2$, distributed as shown:

α	β	A	B	Pr
1	1	1	1	.140
		1	2	.360
		2	1	.360
		2	2	.140
1	2	1	1	.198
		1	2	.302
		2	1	.302
		2	2	.198
2	1	1	1	.189
		1	2	.311
		2	1	.311
		2	2	.189
2	2	1	1	.460
		1	2	.040
		2	1	.040
		2	2	.460

Marginal selectivity here is satisfied trivially: all marginal probabilities are equal 0.5, for all treatments. In the matrix form of the LFT, the column-vector of the above 16 probabilities,

$$(.140, .360, .360, \ldots, .040, .040, .460)^{\top},$$

using $\top$ for transposition, is denoted by P. The LFT problem is defined by the system $MQ=P$, $Q \geq 0$, where the 16×16 Boolean matrix M is shown below: each column of the matrix corresponds to a combination of values for the hypothetical H-variables (shown above the matrix), while each row corresponds to a combination of a treatment with values of the outputs A, B (shown on the left).

			H_{1^α}	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2
			H_{2^α}	1	1	1	1	2	2	2	2	1	1	1	1	2	2	2	2
			H_{1^β}	1	1	2	2	1	1	2	2	1	1	2	2	1	1	2	2
			H_{2^β}	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2
α	β	A	B																
1	1	1	1	1	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0
		1	2	0	0	1	1	0	0	1	1	0	0	0	0	0	0	0	0
		2	1	0	0	0	0	0	0	0	0	1	1	0	0	1	1	0	0
		2	2	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	1
1	2	1	1	1	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0
		1	2	0	1	0	1	0	1	0	1	0	0	0	0	0	0	0	0
		2	1	0	0	0	0	0	0	0	0	1	0	1	0	1	0	1	0
		2	2	0	0	0	0	0	0	0	0	0	1	0	1	0	1	0	1
2	1	1	1	1	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0
		1	2	0	0	1	1	0	0	0	0	0	0	1	1	0	0	0	0
		2	1	0	0	0	0	1	1	0	0	0	0	0	0	1	1	0	0
		2	2	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	1
2	2	1	1	1	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0
		1	2	0	1	0	1	0	0	0	0	0	1	0	1	0	0	0	0
		2	1	0	0	0	0	1	0	1	0	0	0	0	0	1	0	1	0
		2	2	0	0	0	0	0	1	0	1	0	0	0	0	0	1	0	1

The linear programing routine of Mathematica™(using the interior point algorithm) shows that here the linear system (2) has solutions corresponding to the JDC-vector

H_{1^α}	H_{2^α}	H_{1^β}	H_{2^β}	Pr
1	1	1	1	.02708610
1	1	1	2	.00239295
1	1	2	1	.16689300
1	1	2	2	.03358610
1	2	1	1	.00197965
1	2	1	2	.10854100
1	2	2	1	.00204128
1	2	2	2	.15748000

H_{1^α}	H_{2^α}	H_{1^β}	H_{2^β}	Pr
2	1	1	1	.15748000
2	1	1	2	.00204128
2	1	2	1	.10854100
2	1	2	2	.00197965
2	2	1	1	.03358610
2	2	1	2	.16689300
2	2	2	1	.00239295
2	2	2	2	.02708610

The column-vector of these probabilities constitutes $Q > 0$. This proves that in this case we do have $(A, B) \looparrowleft (\alpha, \beta)$.

Let us now change the distributions of (A, B) to the following:

α	β	A	B	Pr
1	1	1	1	.450
		1	2	.050
		2	1	.050
		2	2	.450

α	β	A	B	Pr
1	2	1	1	.105
		1	2	.395
		2	1	.395
		2	2	.105

α	β	A	B	Pr
2	1	1	1	.170
		1	2	.330
		2	1	.330
		2	2	.170

α	β	A	B	Pr
2	2	1	1	.110
		1	2	.390
		2	1	.390
		2	2	.110

Once again, marginal selectivity is satisfied trivially, as all marginal probabilities are 0.5, for all treatments. The linear programing routine of Mathematica$^{\text{TM}}$, however, shows that the linear system (2) has no solutions here. This excludes the existence of a JDC-vector for this situations, ruling out thereby the possibility of $(A, B) \looparrowleft (\alpha, \beta)$.

2.4 Paralells with Quantum Physics

Both the Linear Feasibility Test and the Joint Distribution Criterion on which it is based have their analogues in quantum physics. To appreciate the analogy, however, one has to adopt the interpretation of noncommuting quantum measurements performed on a given component of a quantum-entangled system as mutually exclusive factor levels of the same factor.

In the Einstein-Podolsky-Rosen (EPR) paradigm, several subatomic particles are emitted from a common source in such a way that they remain *entangled* (have highly correlated properties, such as momenta or spins) as they run away from each other. An experiment may consist, e.g., in measuring the spin of electron 1 along one of several axes, $\alpha^1 = \alpha^1_1$, $\alpha^1 = \alpha^1_2$, etc., and (in another location but simultaneously in some frame of reference) measuring the spin of electron 2 along one of several axes, $\alpha^2 = \alpha^2_1$, $\alpha^2 = \alpha^2_2$, etc., and in the same manner for other particles. The outcome A^i of a measurement along any axis on particle $i = 1, \ldots, n$ is a random variable with several possible values, depending on the spin number of the particles (for electrons, there are two possible values, "up" or "down"). The question that arises is: does measurement α^i selectively affect only A^i (even though $A^1, \ldots, A^n$ are not independent)? If the answer is negative, then the measurement of one electron affects the outcome of the measurement of another electron even though no information can be exchanged between two distant events that are simultaneous in some frame of reference. What makes this situation formally identical to the selective influence problems considered above is that the measurements along two different axes, say, α_1^i and α_2^i, are *non-commuting*, i.e., they cannot be performed on the ith particle simultaneously. This makes it possible to consider them as levels of factor α^i.

Below is the table of correspondences between the general language of selective probabilsitic causality and the quantum-mechanical notions used in the analysis of spins of entangled particles:

Selective Probabilistic Causality (general)	Quantum Entanglement Problem (for spins)
observed random output	detected spin value of a given particle
factor/input	spin measurement in a given particle
factor level	setting (axis) of the spin measurement
joint distribution criterion	joint distribution criterion
canonical diagram of selective influences	"classical" explanation (by context-independent local variables)
representation in the form SI_1	probabilistic "classical" explanation
representation in the form SI_2	deterministic "classical" explanation

The results of the simplest entanglement experiment, with $n = 2, k_1 = k_2 = 2$, $m_1 = m_2 = 2$, are described by the estimates of 16 probabilities

$$p\left(A^1, A^2|\alpha^1, \alpha^2\right) = \Pr\left[A^1 = \left\{\begin{array}{c} up \\ down \end{array}\right., A^2 = \left\{\begin{array}{c} up \\ down \end{array}\right. \middle| \alpha^1 = \left\{\begin{array}{c} \alpha_1^1 \\ \alpha_2^1 \end{array}\right., \alpha^2 = \left\{\begin{array}{c} \alpha_1^2 \\ \alpha_2^2 \end{array}\right. \right],$$

with nothing preventing one, of course, from encoding both α_1^1 and α_1^2 by 1 and α_2^1 and α_2^2 by 2. Encoding "down" and "up" spins for A by $\bullet$ and $\circ$, and for B by $\sqcup$ and $\sqcap$, we get

$\phi = (1,1)$	$B_{11} = \sqcup$	$B_{11} = \sqcap$	
$A_{11} = \bullet$	p_{11}	p_{12}	$a_{1\cdot}$
$A_{11} = \circ$	p_{21}	p_{22}	$a_{2\cdot}$
	$b_{\cdot 1}$	$b_{\cdot 2}$	

$\phi = (1,2)$	$B_{12} = \sqcup$	$B_{12} = \sqcap$	
$A_{12} = \bullet$	q_{11}	q_{12}	$a_{1\cdot}$
$A_{12} = \circ$	q_{21}	q_{22}	$a_{2\cdot}$
	$b'_{\cdot 1}$	$b'_{\cdot 2}$	

$\phi = (2,1)$	$B_{21} = \sqcup$	$B_{21} = \sqcap$	
$A_{21} = \bullet$	r_{11}	r_{12}	$a'_{1\cdot}$
$A_{21} = \circ$	r_{21}	r_{22}	$a'_{2\cdot}$
	$b_{\cdot 1}$	$b_{\cdot 2}$	

$\phi = (2,2)$	$B_{22} = \sqcup$	$B_{22} = \sqcap$	
$A_{22} = \bullet$	s_{11}	s_{12}	$a'_{1\cdot}$
$A_{22} = \circ$	s_{21}	s_{22}	$a'_{2\cdot}$
	$b'_{\cdot 1}$	$b'_{\cdot 2}$	

It is known since Arthur Fine's work (*J. Math. Phys.* 23, 1306-1310, 1982) that the existence of the JDC-vector for this situation (interpreted as the existence of a classical explanation for it) is equivalent to the probabilities satisfying the following inequalities:

$$\begin{array}{l} -1 \le p_{11} + r_{11} + s_{11} - q_{11} - a'_{1\cdot} - b_{\cdot 1} \le 0, \\ -1 \le q_{11} + s_{11} + r_{11} - p_{11} - a'_{1\cdot} - b'_{\cdot 1} \le 0, \\ -1 \le r_{11} + p_{11} + q_{11} - s_{11} - a_{1\cdot} - b_{\cdot 1} \le 0, \\ -1 \le s_{11} + q_{11} + p_{11} - r_{11} - a_{1\cdot} - b'_{\cdot 1} \le 0. \end{array} \quad (3)$$

By applying the LFT to the matrices above, these inequalities are shown to be solutions for (2), with

$$P = (p_{11}, p_{12}, \ldots, s_{21}, s_{22})^\top$$

and M the same 16×16 Boolean matrix as in Section 2.3. In fact, using a standard facet enumeration program (e.g., lrs program at http://cgm.cs.mcgill.ca/~avis/C/lrs.html) these inequalities (together with the equalities representing marginal selectivity) can be derived from (2) "mechanically."

The same mechanical derivation can be used for larger entanglement problems. Once such a system of inequalities S is derived, one can use it to prove necessity (or sufficiency) of any other system S' by showing, with the aid of a linear programming algorithm, that S' is redundant when added to S (respectively, S is redundant when added to S'). But given a set of numerical (experimentally estimated or theoretical) probabilities, computing $\mathcal{L}(P)$ is always preferable to dealing with explicit inequalities as their number becomes very large even for moderate-size vectors P. While the set of inequalities (for $n = 2$, $k_1 = k_2 = 2$, $m_1 = m_2 = 2$), assuming that the marginal selectivity equalities hold, number just 8, already for $n = 2$, $k_1 = k_2 = 2$ with $m_1 = m_2 = 3$ (describing, e.g., an EPR experiment with two spin-1 particles, or two spin-$1/2$ ones and inefficient detectors), our computations yield 1080 inequalities equivalent to $\mathcal{L}(P) = 1$, and for $n = 3$, $k_1 = k_2 = k_3 = 2$ and $m_1 = m_2 = m_3 = 2$, corresponding to the Greenberger-Horne-Zeilinger paradigm with three spin-$1/2$ particles, this number is 53792.

The inequalities in (3) can also be derived using the classification distances discussed in Section 2.2. Consider the chain inequalities for the order-distance D_1 obtained by putting $\bullet = \sqcup = 1$, $\circ = \sqcap = 2$, and identifying $\preceq$ with $\leq$:

$$\begin{aligned}
q_{12} = \mathrm{D}_1(H^1_x,H^2_{y'}) &\leq \mathrm{D}_1(H^1_x,H^2_y)+\mathrm{D}_1(H^2_y,H^1_{x'})+\mathrm{D}_1(H^1_{x'},H^2_{y'}) = p_{12}+r_{21}+s_{12},\\
p_{12} = \mathrm{D}_1(H^1_x,H^2_y) &\leq \mathrm{D}_1(H^1_x,H^2_{y'})+\mathrm{D}_1(H^2_{y'},H^1_{x'})+\mathrm{D}_1(H^1_{x'},H^2_y) = q_{12}+s_{21}+r_{12},\\
s_{12} = \mathrm{D}_1(H^1_{x'},H^2_{y'}) &\leq \mathrm{D}_1(H^1_{x'},H^2_y)+\mathrm{D}_1(H^2_y,H^1_x)+\mathrm{D}_1(H^1_x,H^2_{y'}) = r_{12}+p_{21}+q_{12},\\
r_{12} = \mathrm{D}_1(H^1_{x'},H^2_y) &\leq \mathrm{D}_1(H^1_{x'},H^2_{y'})+\mathrm{D}_1(H^2_{y'},H^1_x)+\mathrm{D}_1(H^1_x,H^2_y) = s_{12}+q_{21}+p_{12}.
\end{aligned} \tag{4}$$

Consider also the inequalities for the order-distance D_2 obtained by putting $\bullet = \sqcap = 1$, $\circ = \sqcup = 2$, and identifying $\preceq$ with $\leq$:

$$\begin{aligned}
q_{11} = \mathrm{D}_2(H^1_x,H^2_{y'}) &\leq \mathrm{D}_2(H^1_x,H^2_y)+\mathrm{D}_2(H^2_y,H^1_{x'})+\mathrm{D}_2(H^1_{x'},H^2_{y'}) = p_{11}+r_{22}+s_{11},\\
p_{11} = \mathrm{D}_2(H^1_x,H^2_y) &\leq \mathrm{D}_2(H^1_x,H^2_{y'})+\mathrm{D}_2(H^2_{y'},H^1_{x'})+\mathrm{D}_2(H^1_{x'},H^2_y) = q_{11}+s_{22}+r_{11},\\
s_{11} = \mathrm{D}_2(H^1_{x'},H^2_{y'}) &\leq \mathrm{D}_2(H^1_{x'},H^2_y)+\mathrm{D}_2(H^2_y,H^1_x)+\mathrm{D}_2(H^1_x,H^2_{y'}) = r_{11}+p_{22}+q_{11},\\
r_{11} = \mathrm{D}_2(H^1_{x'},H^2_y) &\leq \mathrm{D}_2(H^1_{x'},H^2_{y'})+\mathrm{D}_2(H^2_{y'},H^1_x)+\mathrm{D}_2(H^1_x,H^2_y) = s_{11}+q_{22}+p_{11}.
\end{aligned} \tag{5}$$

It was shown that

> Each right-hand inequality in (3) is equivalent to the corresponding chain inequality in (4) for the order-distance D_1. Each left-hand inequality in (3) is equivalent to the corresponding chain inequality in (5) for the order-distance D_2.

2.5 Sample-level tests

The set of vectors P for which the system (2) has a solution forms a convex polytope. Recently Clintin Davis-Stober (*J. Math. Psych.*, 53, 1–13, 2009) developed a statistical theory for testing the hypothesis that a vector of probabilities P (not necessarily of the same structure as in LFT) belongs to a convex polytope $\mathcal{P}$ against the hypothesis that it does not. Under certain regularity constraints he derived the asymptotic distribution (a convex mixture of chi-square distributions) for the log maximum likelihood ratio statistic

$$-2 \log \frac{\max_{P \in \mathcal{P}} L(P|N)}{\max_P L(P|N)},$$

where N is the vector of observed absolute frequencies, comprised of the numbers of occurrences of $(l_1, \ldots, l_n; j_1, \ldots, j_n)$ in the case of LFT. The likelihoods $L(P|N)$ were computed using the standard theory of multinomial distributions. Due to this development the statistical aims of this part of the project were deemphasized.

Other approaches readily suggest themselves. One of them is to use the known theory of $L(P|N) / \max_P L(P|N)$ to compute a confidence region of possible probability vectors P for a given empirical vector N. The hypothesis of selective influences is retained or rejected according as this confidence region contains or does not contain a point P that passes LFT. Resampling techniques is another obvious approach, e.g., the permutation test in which the assignment of empirical distributions to different treatments is randomly "reshuffled" so that each distribution generally ends up assigned to a "wrong" treatment. If the proportion of the permuted assignments whose deviation from the LFT polytope does not exceed that of the the observed estimate of P is sufficiently small, the hypothesis of selective influences can be considered supported.

3 Conclusion

Within the framework of this part of the project,

- a general mathematical theory of selective influences was elaborated (which input influences which of probabilistically interdependent random outputs);
- the Joint Distribution Criterion was formulated in complete generality;
- a theory of pseudo-quasi-metrics was constructed to be used to test for selectiveness of influences;
- a Linear Feasibility Test for selective influences with finite-valued random outputs was constructed;
- a formal equivalence of selective influences with the issue of quantum entanglement in physics was established, with non-commuting measurements in quantum physics paralleling the mutually exclusive values of inputs (external factors) in behavioral sciences.

4 Publications in which the grant is acknowledged

4.1 Peer-reviewed.

Ünlü, A., Kiefer, T., & Dzhafarov, E.N. (2009). Fechnerian Scaling in R: The package fechner. Journal of Statistical Software, 31, Issue 6, 1-24. (URL for the paper and software: http://www.jstatsoft.org/v31/i06.)

Kujala, J.V., & Dzhafarov, E.N. (2009). A new definition of well-behaved discrimination functions. Journal of Mathematical Psychology, 53, 593–599.

Kujala, J.V., & Dzhafarov, E.N. (2009). Regular Minimality and Thurstonian-type modeling. Journal of Mathematical Psychology, 53, 486–501.

Dzhafarov, E.N., & Dzhafarov, D.D. (2010). Sorites without vagueness I: Classificatory sorites. Theoria, 76, 4-24.

Dzhafarov, E.N., & Dzhafarov, D.D. (2010). Sorites without vagueness II: Comparative sorites. Theoria, 76, 25-53.

Dzhafarov, E.N. (2010). Dissimilarity Cumulation as a procedure correcting for violations of triangle inequality. Journal of Mathematical Psychology, 54, 284-287.

Dzhafarov, E.N. (2010). Dissimilarity, quasidistance, distance. Journal of Mathematical Psychology, 54, 334-337.

Dzhafarov, E.N., & Perry, L. (2010). Matching by adjustment: if X matches Y, does Y match X? Frontiers in Quatitative Psychology and Measurement. 1:24. doi:10.3389/fpsyg.2010.00024.

Dzhafarov, E.N., & Kujala, J.V. (2010). The Joint Distribution Criterion and the Distance Tests for Selective Probabilistic Causality. Frontiers in Quantitative Psychology and Measurement. 1:151 doi: 10.3389/fpsyg.2010.00151.

Trendtel, M., Ünlü, A., & Dzhafarov, E.N. (2010). Matrices satisfying Regular Minimality. Frontiers in Quantitative Psychology and Measurement. 1:211 doi: 10.3389/fpsyg.2010.00211.

Dzhafarov, E.N. (2011). Mathematical foundations of Universal Fechnerian Scaling. In B. Berglund, G.B. Rossi, J. Townsend, & L. Pendrill (Eds). Measurements With Persons (pp. 185-210). New York: Psychology Press.

Dzhafarov, E.N., Ünlü, A., Trendtel, M., and Colonius, H. (2011). Matrices with a given number of violations of Regular Minimality. Journal of Mathematical Psychology, 55, 240–250.

Dzhafarov, E.N., and Colonius, H. (2011). The Fechnerian idea. American Journal of Psychology, 124, 127-140.

Dzhafarov, D.D., & Dzhafarov, E.N. (2011). The equivalence of two ways of computing distances from dissimilarities for arbitrary sets of stimuli. Journal of Mathematical Psychology, 55, 469-472.

Dzhafarov, E.N. & Kujala, J.V. (2012). Selectivity in probabilistic causality: Where psychology runs into quantum physics. Journal of Mathematical Psychology, 56, 54-63.

Dzhafarov, E.N., & Kujala, J.V. (in press). Order-distance and other metric-like functions on jointly distributed random variables. Proceedings of the American Mathematical Society.

4.2 Not peer-reviewed

Dzhafarov, E.N. (2009). Review of "Sensory Neuroscience: Four Laws of Psychophysics" by Josef J. Zwislocki, Springer, (2009). ix+170 pp., Index. Journal of Mathematical Psychology, 53, 600–602.

Kiefer, T., Ünlü, A., & Dzhafarov, E.N. (2010). The R package fechner for Fechnerian scaling. In H. Locarek-Junge & C. Weihs (Eds.), Studies in Classification, Data Analysis, and Knowledge Organization (pp. 315–322). Berlin: Springer.

Kujala, J.V., & Dzhafarov, E.N. (2010). Using well-behaved Thurstonian-type models to emulate Regular Minimality. In A. Bastianelli & G. Vidotto (Eds.), Fechner Day 2010 (pp. 45–50). Padua, Italy: The International Society for Psychophysics.

Trendtel, M., Ünlü, A., & Dzhafarov, E.N. (2010). With what probability regular minimality can be satisfied by chance? In A. Bastianelli & G. Vidotto (Eds.), Fechner Day 2010 (pp. 63–68). Padua, Italy: The International Society for Psychophysics.

Dzhafarov, E.N., & Paramei, G.V. (2010). Space of facial expressions: Cumulated versus transformed dissimilarities. In A. Bastianelli & G. Vidotto (Eds.), Fechner Day 2010 (pp. 605–610). Padua, Italy: The International Society for Psychophysics.

Dzhafarov, E.N. (2011). On the reverse problem of Fechnerian Scaling. In E.N. Dzhafarov & L. Perry (Eds). Descriptive and Normative Approaches to Human Behavior (pp. 91-122). New Jersey: World Scientific.

Dzhafarov, E.N., & Dzhafarov, D.D. (in press as of 2010). The sorites paradox: A behavioral approach. In J. Valsiner and L. Rudolph (Eds). Mathematical Models for Research on Cultural Dynamics: Qualitative Mathematics for the Social Sciences. Routledge: London.

4.3 Archived

Dzhafarov, E.N., & Kujala, J.V. (2011). Selectivity in probabilistic causality: Drawing arrows from inputs to stochastic outputs. arXiv: 1108.3074.

Dzhafarov, E.N., & Kujala, J.V. (2011). Order-distance and other metric-like functions on jointly distributed random variables. arXiv: 1110.1228.

Dzhafarov, E.N., & Kujala, J.V. (2011). Selectivity in probabilistic causality: Where psychology runs into quantum physics. arXiv: 1110.2388.

5 Presentations in which the grant was acknowledged

5.1 Conferences

Perry, L., & Dzhafarov, E.N. (2009, August). Perceptual discrimination of two-dimensional stimuli: a test of matching regularity. Society for Mathematical Psychology Meeting, Amsterdam, The Netherlands.

Kujala, J.V., & Dzhafarov, E.N. (2009, August). Reconciling Regular Minimality with Thurstonian-type Models. Society for Mathematical Psychology Meeting, Amsterdam, The Netherlands.

Kiefer, T., Unlu, A., & Dzhafarov, E.N. (2009, August). Fechnerian Scaling in R. Society for Mathematical Psychology Meeting, Amsterdam, The Netherlands.

Dzhafarov, E.N., & Kujala, J.V. (2010, August). A Criterion and Tests for Selective Probabilistic Causality. Society for Mathematical Psychology Meeting, Portland, Oregon.

Perry, L., & Dzhafarov, E.N. (2010, August). Matching Regularity for 2D Shapes and Locations. Society for Mathematical Psychology Meeting, Portland, Oregon.

Perry, L., & Dzhafarov, E.N. (2010, September). Matching Regularity for 2D Shapes and Locations. European Mathematical Psychology Group Meeting (Jyvaskyla, Finland).

Dzhafarov, E.N., & Kujala, J.V. (2010, September). The Joint Distribution Criterion and the Distance Tests for Selective Probabilistic Causality. European Mathematical Psychology Group Meeting (Jyvaskyla, Finland).

Kujala, J.V., & Dzhafarov, E.N. (2010, October). Regular Minimality Principle and well-behaved Thurstonian-type models. Meeting of the International Society for Psychophysics (Padua, Italy).

Dzhafarov, E.N., & Paramei, G.V. (2010, October). Space of facial expressions: cumulated versus transformed dissimilarities. Meeting of the International Society for Psychophysics (Padua, Italy).

Trendtel, M., Ünlü, A., & Dzhafarov, E.N. (2010, October). With what probability regular minimality can be satisfied by chance? Meeting of the International Society for Psychophysics (Padua, Italy).

Dzhafarov, E.N. (2010, October). Dissimilarity and Regular Minimality. Purdue Winer Memorial Lectures (West Lafayette, Indiana).

Dzhafarov, E.N., & Kujala, J.V. (2011, July). Joint Distribution Criterion and Linear Feasibility Test for selective influences. Society for Mathematical Psychology Meeting, Medford, Massachusetts.

Dzhafarov, E.N., & Kujala, J.V. (2011, July). Can we do without sample spaces? Society for Mathematical Psychology Meeting, Medford, Massachusetts.

Perry, L., & Dzhafarov, E.N. (2011, July). A test of tri-areal matching regularity for two-dimensional stimuli. Society for Mathematical Psychology Meeting, Medford, Massachusetts.

5.2 Individual invited

Swedish Collegium for Adavanced Studies, Uppsala, Sweden (2009, October).
Department of Philosophy, Uppsala University, Sweden (2009, November).
Department of Statistics, University of Dortmund, Germany (2009, December).
Department of Statistics, Purdue University (2011, March).
Department of Psychology, Indiana University (2011, October).